Modeling the Csr System and BarA/UvrY Two Component System (TCS) in response to stress

By

Isabella Joseph

B.S. Chemical Engineering Honors

Undergraduate Engineering Honors Thesis

May 2021

Adviser: Dr. Lydia Contreras

Second Reader: Dr. D'Arcy Randall

McKetta Department of Chemical Engineering

Abstract

The Carbon Storage Regulatory System (Csr) in *Escherichia coli* is a model post-transcriptional regulatory network that includes four components: CsrA protein, CsrB and CsrC noncoding RNAs, and CsrD protein. This system is important to study as it has a wide reach in controlling other regulatory networks throughout *E coli* and is necessary for stress responses. The Carbon Storage Regulatory System responds to external stimuli through the BarA/UvrY Two Component System (TCS). The BarA/UvrY TCS is composed of a histidine kinase, BarA and a response regulator, UvrY, and are primarily connected to the Csr system, as the two proteins serve as a transcriptional activator of CsrB and CsrC. The goal for this project is to understand and computationally model how the Csr and TCS system interact in response to extracellular metabolites. Results of the computational model developed show that incorporating the TCS system allows for quicker response times and could potentially account for the differences in Csr system component concentrations observed in literature. Experimental validation of model with *in vivo* results is required to verify the accuracy of the model. This project is one of the first that works on computationally linking the external metabolite stimuli to the Csr system and could pave the way for advances in the engineering of bacterial genetic circuitry.

Acknowledgements

This project would not have been possible without the help and guidance of many people. Firstly, my sincere thanks to my Principal Investigator, Dr. Lydia Contreras. Thank you for instilling in me the desire to pursue scientific discovery and giving me the platform to conduct research. Huge thanks to Trevor Simmons for helping bring this project to fruition and letting me work on it and always being there to help. Thank you, Dr. Abigail Leistra, for your mentorship and guidance throughout my undergraduate research career. Thanks to every member of the Contreras Biomolecular Lab who have helped me in some way or another during the past four years. Sincere thanks to my Second Reader, Dr. D'Arcy Randall for your feedback on my drafts and helping me become a better writer. Finally, thanks to my parents, brother and roommates for their constant support and encouragement. I am incredibly grateful.

Table of Contents

List of Tables

List of Figures

List of Equations

Introduction

"Bacterial adaptation requires large-scale regulation of gene expression." (Esquerré et al., 2016). Gene expression can be defined as the process of turning on a gene to produce RNA and protein. In order to adapt to changing environmental conditions, organisms have developed various mechanisms of gene expression regulation that allows them to modify their physiology and metabolism. These mechanisms control when a gene is expressed in order to synthesize a protein, to vary the amount of protein made, and to determine when to stop making the protein. Some examples of varying environmental conditions include nutrient deprivation, heat and cold stress and UV radiation. These stress survival behaviors play critical roles in how bacteria respond to antibiotics, live inside a human host, or produce useful chemical products (Winfield & Groisman, 2003).

Escherichia coli are bacteria often found in the gastrointestinal tract of mammals. Due to the highly variable and complex environment of the intestine, *E. coli* often experiences varying growth conditions. In order to cope with the changing conditions in the intestine *E. coli* has developed a variety of stress response mechanisms. Various global regulatory networks found in *E. coli* link physiological and metabolic responses by controlling the functional expression of relevant sets of genes.

The goal for this project is to model BarA/UvrY TCS to investigate how it dynamically responds to extracellular metabolites. This model will give us further insight into how the Csr and TCS system interact to regulate phenotypes like motility, biofilm formation and glycogen metabolism. In this report, I discuss the computational model developed from general material balance principles, the results of the model and the experimental validation process required to test

and retune the model. Development of this model will allow us and other users to more accurately predict the activity of the Csr and TCS systems in response to common external stresses. Building this model will give us the opportunity to build the first model that connects external stimuli to the Csr system.

The Carbon Storage Regulatory System

The Carbon Storage System (Csr) system is a model post-transcriptional regulatory network in *E. coli* that controls a wide variety of physiological adaptative mechanisms. On a phenotype-level, the CsrA protein impacts important cellular behaviors such as glucose storage and processing, synthesis of amino acids, biofilm formation, and motility (Massé, Salvail, Desnoyers & Arguin, 2007).

The Csr system comprises four molecular components: CsrA, CsrB, CsrC and CsrD (Figure 1). The central component of the Csr system is the CsrA, a global regulator protein that binds RNA. It affects the stability and translation of at least 30, but potentially hundreds of mRNAs (Sowa et al., 2017). CsrB and CsrC are two small non-coding RNAs (sRNAs) that negatively regulate CsrA by sequestration. CsrB and CsrC act like a sponge, binding and "capturing" CsrA, which impedes CsrA's ability to bind and regulate target mRNAs (Vakulskas, Potts, Babitzke, Ahmer & Romeo, 2015). CsrB can bind up to 9 copies of the CsrA dimer (Gudapaty, Suzuki, Wang, Babitzke & Romeo, 2001). CsrD is a second protein, which drives the degradation of CsrB and CsrC via RNAse E (Vakulskas, Potts, Babitzke, Ahmer & Romeo, 2015), thus positively regulating CsrA. When CsrA binds to its target mRNA, translation is inhibited as ribosome binding is prevented (Figure 1). This translation inhibition can result in mRNA destabilization. When CsrA is sequestered by CsrB, the ribosome is able to bind to mRNA and translation occurs.

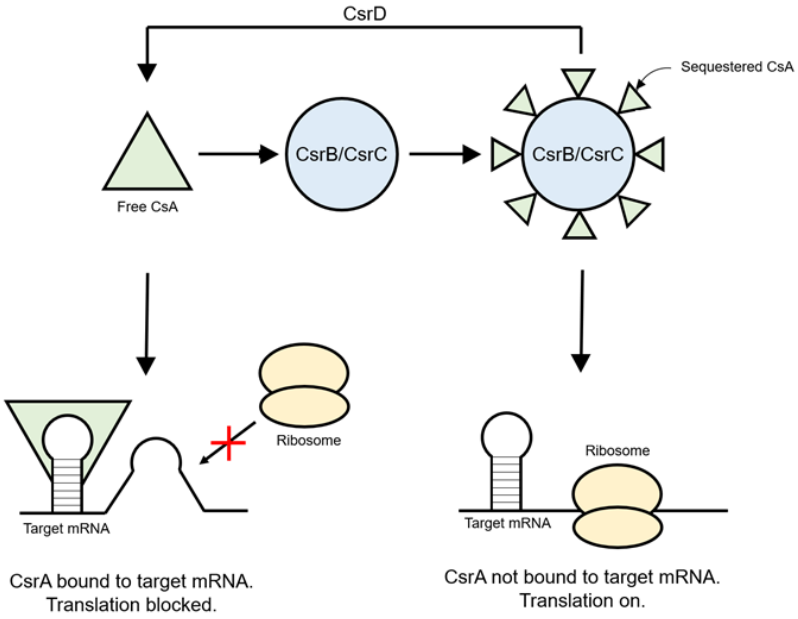

Figure 1: Csr system in *E.coli.*

The Csr system can regulate genome-wide mRNA stability and transcription, thus affecting gene expression in *E. coli* (Esquerré et al., 2016). Approximately 800 mRNA have been discovered across multiple environmental conditions as potentially interacting with CsrA, accounting for about 20% of the *E.coli* genome (Leistra et al., 2018).

The BarA/UvrY Two Component System

Csr sRNA levels are partly regulated by a conserved two-component signal transduction system (TCS) known as BarA-UvrY, one of the many TCS systems in *E.coli* (Figure 2). The BarA/UvrY TCS consists of the membrane-bound associated sensor kinase, BarA and the cytoplasmic response regulator UvrY. The BarA/UvrY two component system plays an important role in central carbon metabolism through transcriptional activation of the *csrB* and *csrC* genes (Mackie, A., 2010). BarA responds to carboxylate-containing metabolites such as acetate and formate. BarA catalyzes the transphosphorylation of its response regulator UvrY, which then

activates transcription from the CsrB and CsrC promoters. glgC-GFP refers to the target mRNA identified in Figure 1. The BarA-UvrY TCS is shown to be advantageous in bacterium's ability to adapt and survive (Pernestig, 2002).

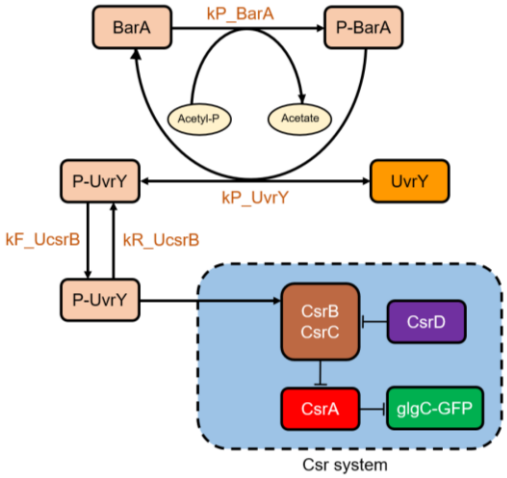

Figure 2: Illustration of the BarA/UvrY TCS interaction with the Csr system via CsrB and CsrC sRNAs transcriptional activation. The abbreviations in the beige font correspond to the nomenclature of the rates described in the "Stability of Constants" section.

CsrA is shown to have a positive effect on BarA/UvrY TCS and is required for proper UvrY expression and activation of BarA kinase activity (Camacho et al., 2014). Acetate and formate act as physiological signals that activate BarA, resulting in transphosphorylation of UvrY. BarA is autophosphorylated and UvrY is phosphorylated by BarA, which then transcriptionally activates CsrB and CsrC expression via binding at the respective promoter regions. CsrB and CsrC sequester the CsrA protein, preventing CsrA from binding to its target mRNAs. CsrA controls the switching of BarA from its phosphatase to its kinase activity and also positively affects uvrY expression (Camacho et al., 2014).

A computational model of the Csr system and the BarA/UvrY TCS system was developed using differential equations. These equations model the concentrations of the different components over time. We expect this model to predict the activity of Csr and TCS and their response to

external stresses more accurately than existing models. Experimental validation of the model will then be conducted in order to ensure accuracy *in vivo*. I was unable to conduct experiments due to COVID-19 restrictions, but I have laid out an experimental plan for the experimental validation process, which can be followed to determine the values for the constants and parameters inputted into the computational model.

Methods

Computational Work

A computational model of the Csr System and BarA/UvrY Two Component System (TCS) was built using existing differential equations derived from general material balance principles. The general balance equation states that the total or component mass or energy of any system can be modeled using Equation 1 shown below. The equation defines the amount of buildup in a system by determining the amount entering and being generated through a reaction and subtracting the amount leaving or consumed. The accumulation of the Csr system and BarA/UvrY TCS components were determined using these general material balance principles.

$$Accumulation = In - Out + Generation - Consumption$$

Equation 1: General Material Balance Equation

The script consists of rate equations that model the concentrations of each component of the Csr and TCS systems, further modified to capture the effect of the TCS system on the Csr system. The input of the model includes an initial concentration of acetate, along with production rates and initial concentrations of the Csr + TCS components. The script consists of differential equations with initial conditions from the inputs. A python built in software is used to solve the differential equations using the initial conditions. The results include the concentrations of all the Csr and TCS components over time, along with the concentrations of external stimuli over time. In the future, this model will then be validated and improved using experiment results.

I determined the differential equations system and parameters following previous computational models of the Csr system as a reference and modified them to include the TCS system (Adamson et al. 2013). The production rate constants were determined by manipulating

constants found in literature. The differential equations used for modeling component concentrations of Csr-mRNA, CsrB-CsrA-mRNA, and the Full Cascade are shown below.

CsrA-mRNA

$$\frac{d[GFP]}{dt} = \alpha_G[m] - (\beta_G + \beta_{dil})[GFP] \quad \textbf{(2)}$$

$$\frac{d[m]}{dt} = \alpha_m - (\beta_m + \beta_{dil})[m] - k_1[A][m] + (k_{-1} + \beta_{AAM})[Am] \quad \textbf{(3)}$$

$$\frac{d[Am]}{dt} = k_1[A][m] - (k_{-1} + \beta_{mAm} + \beta_{AAm} + \beta_{dil})[Am] \quad \textbf{(4)}$$

$$\frac{d[A]}{dt} = \alpha_A - (\beta_A + \beta_{dil})[A] - k_{-1}[A][m] + (k_{-1} + \beta_{mAm})[Am] \quad \textbf{(5)}$$

CsrB-CsrA-mRNA

$$\frac{d[GFP]}{dt} = \alpha_G[m] - (\beta_G + \beta_{dil})[GFP] \quad \textbf{(6)}$$

$$\frac{d[m]}{dt} = \alpha_m - (\beta_m + \beta_{dil})[m] - k_1[A][m] + (k_{-1} + \beta_{AAM})[Am] \quad \textbf{(7)}$$

$$\frac{d[Am]}{dt} = k_1[A][m] - (k_{-1} + \beta_{mAm} + \beta_{AAm} + \beta_{dil})[Am] \quad \textbf{(8)}$$

$$\frac{d[A]}{dt} = \alpha_A - (\beta_A + \beta_{dil})[A] - k_{-1}[A][m] + (k_{-1} + \beta_{mAm})[Am] - k_2[A][B] + (k_{-2} + \beta_{BAB})[AB] \quad \textbf{(9)}$$

$$\boldsymbol{\frac{d[B]}{dt}} = \alpha_B - (\beta_B + \beta_{dil})[B] - k_2[A][B] + (k_{-2} + \beta_{AAB})[AB] \quad \textbf{(10)}$$

$$\boldsymbol{\frac{d[AB]}{dt}} = \boldsymbol{k_2[A][B] - (k_{-2} + \beta_{BAB} + \beta_{AAB} + \beta_{dil})[AB]} \quad \textbf{(11)}$$

Full Cascade

$$\frac{d[GFP]}{dt} = \alpha_G[m] - (\beta_G + \beta_{dil})[GFP] \quad \textbf{(12)}$$

$$\frac{d[m]}{dt} = \alpha_m - (\beta_m + \beta_{dil})[m] - k_1[A][m] + (k_{-1} + \beta_{AAM})[Am] \quad \textbf{(13)}$$

$$\frac{d[Am]}{dt} = k_1[A][m] - (k_{-1} + \beta_{mAm} + \beta_{AAm} + \beta_{dil})[Am] \quad \textbf{(14)}$$

$$\frac{d[A]}{dt} = \alpha_A - (\beta_A + \beta_{dil})[A] - k_{-1}[A][m] + (k_{-1} + \beta_{mAm})[Am] - k_2[A][B] + (k_{-2} + \beta_{BAB})[AB] \quad \textbf{(15)}$$

$$\frac{\boldsymbol{d[B]}}{\boldsymbol{dt}} = \alpha_B - (\omega[D] + \beta_B + \beta_{dil})[B] - k_2[A][B] + (k_{-2} + \beta_{AAB})[AB] \quad \textbf{(16)}$$

$$\frac{\boldsymbol{d[AB]}}{\boldsymbol{dt}} = \boldsymbol{k_2}[\boldsymbol{A}][\boldsymbol{B}] - (\boldsymbol{k_{-2}} + \boldsymbol{\beta_{BAB}} + \boldsymbol{\beta_{AAB}} + \boldsymbol{\beta_{dil}})[\boldsymbol{AB}] \quad \textbf{(17)}$$

$$\frac{\boldsymbol{d[D]}}{\boldsymbol{dt}} = \boldsymbol{\alpha_D} - (\boldsymbol{\beta_D} + \boldsymbol{\beta_{dil}})[\boldsymbol{D}] \quad \textbf{(18)}$$

Equations 2 – 18: Differential equations used to model the CsrA regulatory cascade. (Adamson & Lim, 2013)

Table 1: Variables and constants used in the differential equations (Adamson & Lim, 2013)

Key	
$[GFP]$	Concentration of GlgC-GFP
$[m]$	Concentration of free glgC-gfp mRNA
$[Am]$	Concentration of CsrA-glgC mRNA complex
$[A]$	Concentration of free CsrA
$[B]$	Concentration of free CsrB
$[D]$	Concentration of free CsrD
AB	CsrA– CsrB complex

α_A	Production rate of CsrA
α_B	Production rate of CsrB
α_D	Production rate of CsrD
α_G	rate constant for GFP production
α_m	Production rate of free glgC-gfp mRNA
β_A	Rate constant for active degradation of CsrA
β_B	Rate constant for active degradation of CsrB
$\beta_{\boldsymbol{D}}$	Rate constant for active degradation of CsrD
β_G	Rate constant for active degradation of GFP
β_M	Rate constant for active degradation of mRNA
β_{dil}	Rate constant for passive dilution
β_{AAm}	Rate constant for active degradation of CsrA within the CsrA-glgC mRNA complex
β_{mAm}	Rate constant for degradation of the glgC-gfp mRNA within the CsrA-glgC mRNA complex
β_{AAB}	Rate constant for active degradation of CsrA from the CsrA–CsrB complex
β_{BAB}	Rate constant for active degradation for bound CsrB
k_1	Rate constant for association of the CsrA-glgC mRNA complex
k_{-1}	Rate constant for dissociation of the CsrA-glgC mRNA complex
k_2	Rate constant for association for the CsrA–CsrB complex
k_{-2}	Rate constant for dissociation for the CsrA–CsrB complex
ω	Catalytic efficiency of CsrD-mediated degradation

I developed and refined code using Python to model the Csr with TCS cascade, CsrB activation by TCS and full Csr Cascade. The model predicts the concentration of each component of the Csr and TCS system over time based on various initial conditions, parameters and differential equations determined from literature (Adamson & Lim, 2013; Gonzalez Chavez, Alvarez, Romeo & Georgellis, 2010; Singh, Ekka & Kumaran, 2012; Yamamoto et al., 2005). The computational model will be validated using experimental data in the future. We expect the experimental data to show that this model accurately reflects the native activity of this cascade.

Csr with TCS cascade

The Adamson and Lim paper uses a completely synthetic Csr cascade model and the response of the system is determined by turning ON each Csr cascade component. By contrast, the addition of the BarA/UvrY TCS equations below connects the Csr response to external stimuli, such as acetate in this case.

$$\frac{d[BarA]}{dt} = \alpha_{BarA} - (\beta_{BarA} + \beta_{dil})[BarA] - k_{P-BarA}[BarA][acetate] \quad \textbf{(19)}$$

$$\frac{d[P-BarA]}{dt} = k_{P-BarA}[BarA][acetate] - (\beta_{P-BarA} + \beta_{dil})[P-BarA] - k_{P-UvrY}[UvrY][P-BarA] \quad \textbf{(20)}$$

$$\frac{d[UvrY]}{dt} = \alpha_{UvrY} - (\beta_{UvrY} + \beta_{dil})[UvrY] - k_{P-UvrY}[UvrY][P-BarA] \quad \textbf{(21)}$$

$$\frac{d[P-UvrY]}{dt} = k_{P-UvrY}[UvrY][P-BarA] - (\beta_{P-UvrY} + \beta_{dil})[P-UvrY] - k_{F-UcsrB}[P-UvrY] + k_{R-UcsrB}[P-UvrY-bound] \quad \textbf{(22)}$$

$$\frac{d[P-UvrY-bound]}{dt} = k_{F-UcsB}[P-UvrY] - k_{R-UcsB}[P-UvrY-bound] \quad \textbf{(23)}$$

Equations 19-23 Modified Rate Equations

Table 2 Variables and constants used in the modified differential equations

Key	
$[BarA]$	Concentration of BarA
$[UvrY]$	Concentration of UvrY
$[acetate]$	Concentration of acetate
$[P-BarA]$	Concentration of phosphorylated BarA
$[P-UvrY]$	Concentration of phosphorylated UvrY
$[P-UvrY-bound]$	Concentration of phosphorylated bound UvrY
α_{BarA}	Production rate of BarA
α_{UvrY}	Production rate of UvrY
β_{BarA}	Rate constant for active degradation of BarA
β_{P-BarA}	Rate constant for active degradation of phosphorylated BarA
β_{UvrY}	Rate constant for active degradation of UvrY
β_{P-UvrY}	Rate constant for active degradation of phosphorylated UvrY
β_{dil}	Rate constant for passive dilution
k_{P-BarA}	Rate of BarA phosphorylation by acetate
k_{P-UvrY}	Rate of UvrY phosphorylation by by phosphorylated BarA
k_{csrB}	Upregulated rate of csrBC transcription by phosphorylated UvrY binding
$k_{F-UcsrB}$	Rate of binding of phosphorylated UvrY to csrBC promoter
$k_{R-UcsrB}$	Rate of dissociation of phosphorylated UvrY and csrBC promoter

The code developed integrates the synthetic Csr cascade model with the BarA/UvrY TCS differential equations to result in a model that we hypothesize more accurately predicts the activity of Csr and TCS components in response to stress. Initial conditions for the concentration of each Csr component was entered based on literature values. The initial concentrations of the BarA/UvrY components were initially set to 0. The model predicts the concentration of the Csr+TCS components over time and also compares the difference in concentrations for both the Csr only and the Csr+TCS system.

A perturbation analysis was conducted in order to test the stability of the rate constants used in the model. I varied the concentrations of these constants and observed how they affect response time. The response time is the time required to reduce [CsrA] to half its initial value. We used this metric for the response time as it has been previously used for Csr modeling in various papers (Adamson & Lim, 2013). The results of the perturbation analysis will help determine the constants that need to be experimentally validated first.

Experimental Plan

My first experimental goal for the future is to validate the *in vivo* synthetic CsrB-CsrA cascade demonstrated by Adamson and Lim. I am particularly interested in replicating Figure 3 shown below, which shows the comparison of the synthetic and natively expressed Csr cascade, in which CsrB expression is activated. We hypothesize that the differences in response between the purely synthetic and native systems (shown in Figure 3), may be due to the native activity of the TCS system.

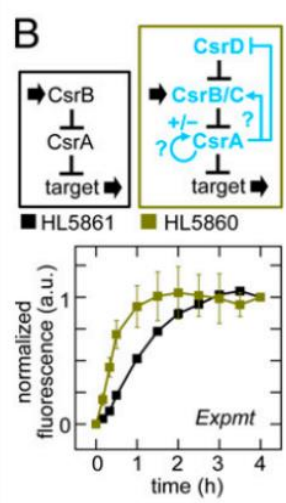

Figure 3: *In vivo* synthetic CsrB-CsrA cascade demonstrated by Adamson and Lim (2013).
The gold plot shows the cascade with native csrA, csrB, csrC and csrD and synthetic csrB, whereas the black plot shows the synthetic cascade

To validate the Csr + TCS model with *in vivo* results, we propose a three plasmid system that links the native Csr + TCS response to a fluorescence output, detailed in Figure 4.

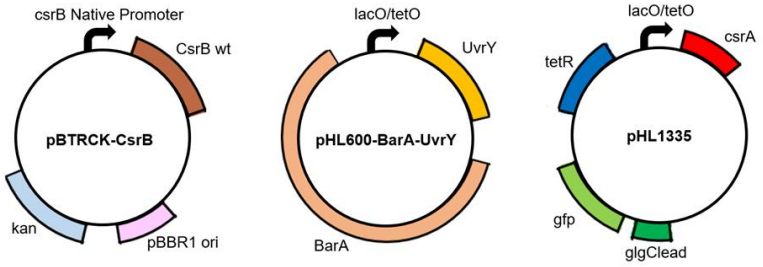

Figure 4: Proposed three plasmid system.
The three plasmid system consists of pBTRCK-CsrB with native csrBC promoter, pHL600-BarA-UvrY with both BarA and UvrY under PLlacO promoter and pHL1335 plasmid with CsrA and glgC-GFP reporter.

Firstly, we will express csrBC and its native promoter on a low copy plasmid to mimic native levels of csrB expression, which matches the current parameters of our model. Next, we will express BarA and UvrY on a single plasmid under the inducible PLLacO promoter, such that we can control the expression of both proteins and ensure they are at similar levels in the cell (Lutz, 1997). Finally, we will express CsrA and the fluorescent reporter on a third plasmid. We will correlate GFP fluorescence to concentration estimates using previously determined Hill Transfer

Functions[11] that correlate fluorescence to GFP concentration and then measure the concentration of CsrA using Western Blot. Additionally, we will perform a phos-tagged Western Blot to measure concentrations of phosphorylated and unphosphorylated UvrY. Based on the results of the experiments, we will refine the constants using the streamlined approach from the perturbation analysis detailed later.

Results and Discussion

Computational Work

Output Concentrations

The code developed models the varying concentrations of Csr+TCS components over time after the activation of the BarA/UvrY signaling cascade. Figure 5 shows the output of the model based on parameters used (Table 4). The parameters were determined from estimations used in literature and they can be modified to observe their effect on the output of the model.

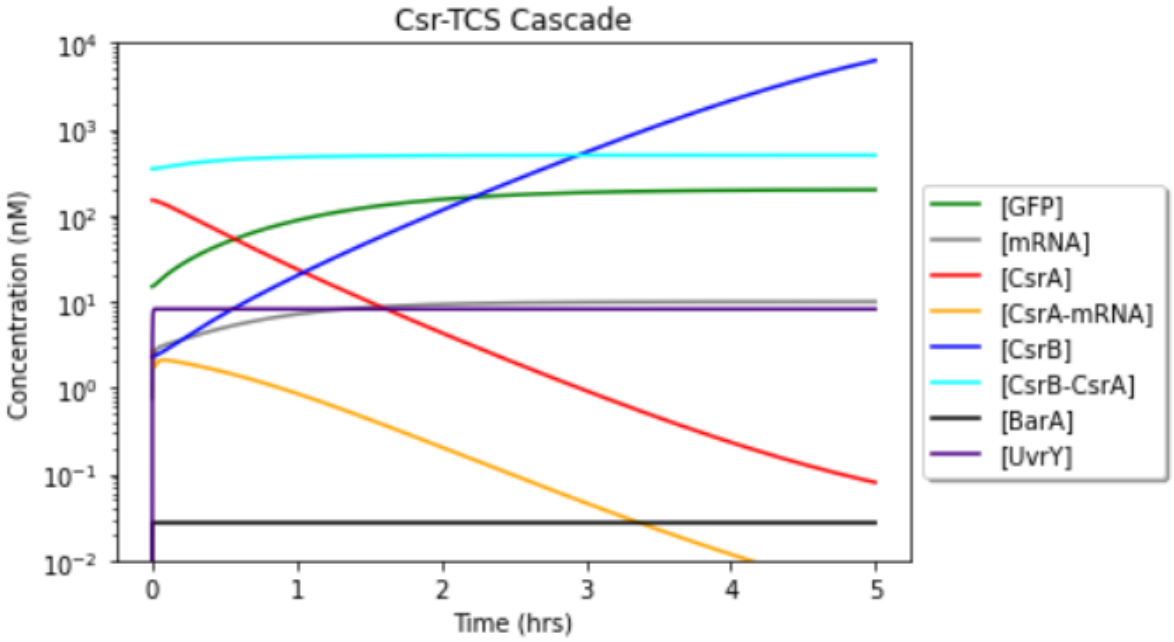

Figure 5: Csr+TCS model component concentrations over time

Table 3: Parameters used in the Csr+TCS model

Csr system parameters			**BarA/UvrY TCS parameters**		
Constant	*Value*	*Units*	*Constant*	*Value*	*Units*
β_G	0	s^{-1}	α_{BarA}	0.5	nM s^{-1}
β_A	0	s^{-1}	β_{BarA}	0	s^{-1}
β_{AAm}	0	s^{-1}	α_{UvrY}	0.5	nM s^{-1}
β_{dil}	0.0004	s^{-1}	k_{P-UvrY}	0.00741	nM^{-1} s^{-1}
α_G	0.008	s^{-1}	β_{UvrY}	0	s^{-1}
β_M	0.006	s^{-1}	β_{P-UvrY}	0.00331	nM^{-1} s^{-1}
α_m	0.064	nM s^{-1}	$k_{F-UcsrB}$	0.0002	nM s^{-1}
k_1	0.0001127	nM s^{-1}	$k_{R-UcsrB}$	0.00013	nM s^{-1}
k_{-1}	0.0051	s^{-1}	k_{csrB}	0.02583	s^{-1}
β_{mAm}	0.02	s^{-1}	$[acetate]$	0.0000028	nM
α_A	0.2	nM s^{-1}			
β_B	0.0004	s^{-1}			
β_{BAB}	0.0004	s^{-1}			
β_{AAB}	0	s^{-1}			
k_2	0.327	nM^{-1} s^{-1}			
k_{-2}	0.327	s^{-1}			
α_B	4	nM s^{-1}			
β_D	0	s^{-1}			
ω	0.008				

From Figure 5, we can see that the concentration of CsrB increases over time and consequently, the concentration of free CsrA decreases. CsrB sequesters CsrA resulting in the decrease of CsrA and an initial increase in the concentration of CsrA–CsrB complex. As CsrA unbounds from the mRNA, the concentration of free *glgC-gfp* mRNA also increases. The concentration of GlgC-GFP increases initially and stabilizes. The response time for turning OFF the Csr system was calculated to be 69.55 minutes. As mentioned above, the response time refers to the time taken to reduce the concentration of CsrA to half its initial value. We see that the BarA/UvrY TCS components stay constant over time. The concentration of BarA and UvrY experience an initial increase in concentration as a result of the activation of the BarA/UvrY signaling cascade, after which both concentrations remain constant.

Comparison of Csr+TCS model and Csr only model

In order to compare the effect of the BarA/UvrY TCS system on the Csr system, a plot of the concentration of Csr components with and without the effect of the TCS system was created (Figure 6). The graph compares the concentrations of free CsrA, free CsrB and CsrA-glgC mRNA complex.

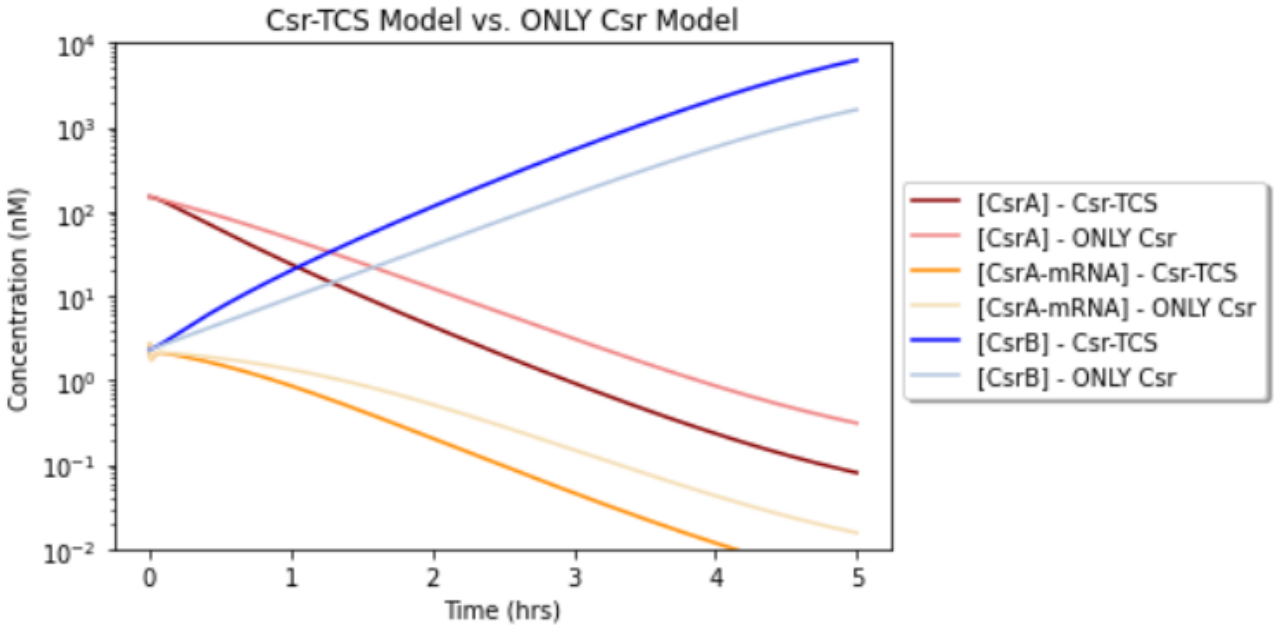

Figure 6: Csr+TCS model component concentrations compared to the Csr only model

We can see that the concentrations of each component follow the same general trend; however, the concentrations at a given time are significantly different, and the difference in concentrations increases over time. For example, the concentration of CsrA-glgC mRNA complex using only the Csr model takes about an hour more to reach the same concentration of CsrA-glgC mRNA complex with the Csr+TCS model.

Overall, we expect that incorporating the BarA/UvrY TCS system accounts for the discrepancy in concentrations observed in Figure 3, as the model used by Adamson and Lim uses a Csr only model. In order to refine the Csr+TCS model built, experimental validation is required. The parameters used in the model will be more accurately determined based on experimental data. $k_{F-UcsrB}$ will be the first parameter to be tested and validated, as the perturbation analysis demonstrated $k_{F-UcsrB}$to have the greatest impact on model response time (shown below).

Perturbation analysis

Before we attempt the experimental validation, we first needed to determine which constants most affect the response time. This knowledge enables us to quicky assess which constants need to be modified. As part of this process, I tested the stability of the phosphorylation constants used in the ODEs. The nomenclature used for each constant is shown in Figure 2 and the ODEs used are detailed in the methods section.

Response time serves as a metric to compare the model across varying conditions, as well as a means to compare it to the *in vivo* system. Figures 7, 8 and 9 show the response times of the model. kP_BarA is the rate of BarA phosphorylation by acetate and kP_UvrY is phosphorylation rate of UvrY by BarA-P. Response time of the model remains constant after varying either kP_BarA or kP_UvrY across several orders of magnitude, as seen in Figures 7 and 8.

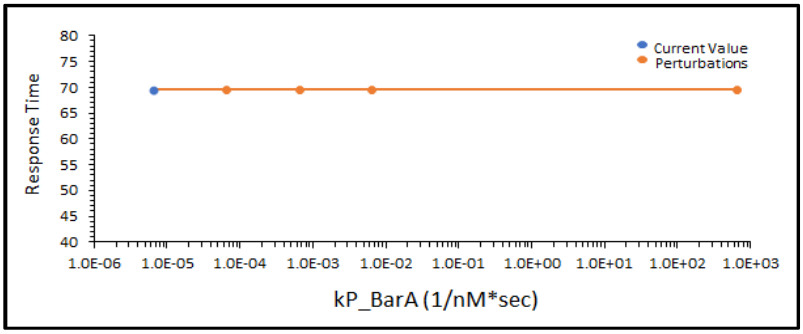

Figure 7: kP_BarA (Rate of BarA Phosphorylation by acetate)

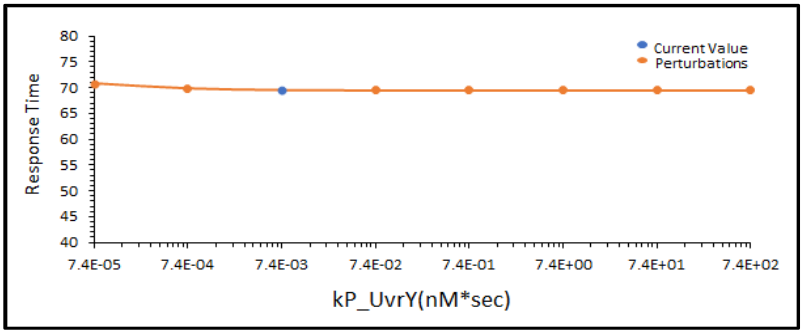

Figure 8: kP_UvrY (Phosphorylation rate of UvrY by BarA-P)

As the response time did not vary drastically with the different values of phosphorylation constants, we can conclude the constants may not significantly affect the response time of the model. However, small perturbations to the binding rate of phosphorylated UvrY to the csrBC promoter (kF_UcsrB) leads to drastic changes in response time of the model (Figure 9).

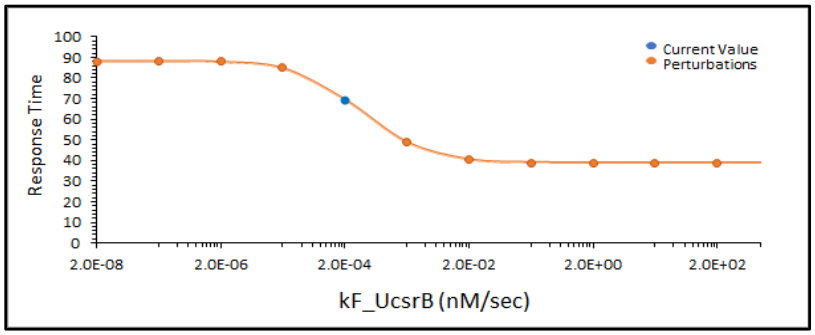

Figure 9: kF_UcsrB (Binding rate of UvrY-P to csrBC promoter)

We notice that the response time is constant after a certain threshold; this effect occurs because the parameter is currently set to an initial UvrY concentration. If the current model requires additional iterations to match the *in vivo* system, we will focus first on refining the kF_UcsrB constant, because these results show it to have the greatest impact on model response time.

Experimental Validation

Based on the experimental results, the constants used in the computational model will be refined using the constant stability method described previously. The first rate constant to be tested will be $k_{F-UcsrB}$. The model currently uses a value of 0.0002 nM s^{-1} for $k_{F-UcsrB}$. If a discrepancy is observed between the experimental results and the output of the model, these constants will be revalidated using experimental data. The experimental validation process using mock data is shown in Figure 10.

Upon validation of these results, we will refine our current TCS + Csr model to match the *in vivo* results. Development of this model will allow us and other users to more accurately predict the activity of the Csr and TCS systems in response to common external stresses difficult to replicate in a laboratory setting, such as stochastic or oscillatory metabolite stimuli.

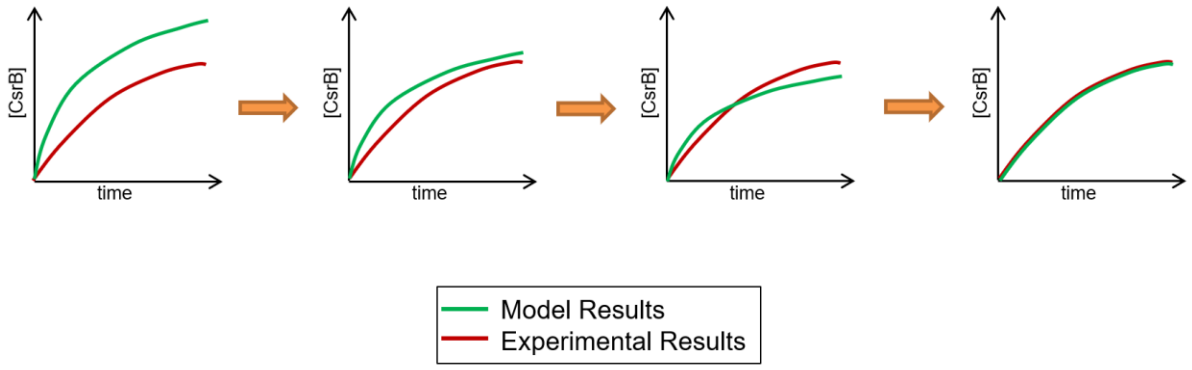

Figure 10: Experimental Validation Process

Conclusion and Future Work

The goal of this project was to computationally model BarA/UvrY TCS to investigate how it dynamically responds to extracellular metabolites and show that the TCS system and its ability to rapidly activate CsrB expression accounts for the differences observed between the synthetic and native cascades in literature. Development of this model allows us and other users to more accurately predict the activity of the Csr and TCS systems in response to common external stresses.

In this study, I developed a computational model by using general material balance principles to model the concentrations of both Csr system components and the BarA/UvrY system components. The Csr-TCS cascade model allows us to visualize the change in concentrations of GFP, mRNA, CsA, CsrA-mRNA, CsrB, CsrB-CsrA, BarA and UvrY over time. A script was also developed to compare how the Csr+TCS model compares to the Csr only model. The results show that the response time is significantly lower when the TCS system is incorporated with the Csr system. The TCS system and its ability to rapidly activate CsrB expression accounts for the differences observed between the synthetic and native cascades.

Future work on this project involves the validation of the computational model with *in vivo* results. This experimental validation will suggest ways to modify the constants and parameters used in the model to more accurately predict the Csr+TCS response to stress. Due to COVID-19 safety precautions, I was unable to enter the lab and conduct experiments for the experimental validation process. This report shows the design of the necessary plasmids and primers and sets out the experimental plan to allow for a smooth data collection.

Developing an accurate model of how the TCS responds to stress will provide further insight into how Csr systems dynamically respond to external stresses and signals. We expect that incorporating the BarA/UvrY TCS system in the Csr system accounts for the discrepancy in

concentrations observed by Adamson and Lim. This project is one of the first that connects external stimuli to the Csr system and could pave the way for potential engineering of bacterial genetic circuitry as the more we understand the system the more we can engineer it.

References

Adamson, D., & Lim, H. (2013). Rapid and robust signaling in the CsrA cascade via RNA-protein interactions and feedback regulation. Proceedings Of The National Academy Of Sciences, 110(32), 13120-13125. doi: 10.1073/pnas.1308476110

Camacho, M., Alvarez, A., Gonzalez Chavez, R., Romeo, T., Merino, E., & Georgellis, D. (2014). Effects of the Global Regulator CsrA on the BarA/UvrY Two-Component Signaling System. Journal Of Bacteriology, 197(5), 983-991. doi: 10.1128/jb.02325-14

Esquerré, T., Bouvier, M., Turlan, C., Carpousis, A., Girbal, L., & Cocaign-Bousquet, M. (2016). The Csr system regulates genome-wide mRNA stability and transcription and thus gene expression in Escherichia coli. Scientific Reports, 6(1). doi: 10.1038/srep25057

Gudapaty, S., Suzuki, K., Wang, X., Babitzke, P., & Romeo, T. (2001). Regulatory Interactions of Csr Components: the RNA Binding Protein CsrA Activates csrB Transcription inEscherichia coli. Journal Of Bacteriology, 183(20), 6017-6027. doi: 10.1128/jb.183.20.6017-6027.2001

Gonzalez Chavez, R., Alvarez, A., Romeo, T., & Georgellis, D. (2010). The Physiological Stimulus for the BarA Sensor Kinase. Journal Of Bacteriology, 192(7), 2009-2012. doi: 10.1128/jb.01685-09

Leistra, A., Gelderman, G., Sowa, S., Moon-Walker, A., Salis, H., & Contreras, L. (2018). A Canonical Biophysical Model of the CsrA Global Regulator Suggests Flexible Regulator-Target Interactions. Scientific Reports, 8(1). doi: 10.1038/s41598-018-27474-2

Lutz, R. (1997). Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. Nucleic Acids Research, 25(6), 1203-1210. doi: 10.1093/nar/25.6.1203

Mackie, A., 2010. Escherichia Coli K-12 Substr. MG1655 Bara Uvry Two-Component Signal Transduction System. [online] Biocyc.org. Available at https://biocyc.org/ECOLI/new-image?type=PATHWAY&object=PWY0-1501

Massé, E., Salvail, H., Desnoyers, G., & Arguin, M. (2007). Small RNAs controlling iron metabolism. Current Opinion In Microbiology, 10(2), 140-145. doi: 10.1016/j.mib.2007.03.013

Pernestig, A. (2002). The study of the Escherichia coli BarA-UvrY two-component system and its ability to sense the environment. Stockholm.

Singh, V., Ekka, M., & Kumaran, S. (2012). Second Monomer Binding Is the Rate-Limiting Step in the Formation of the Dimeric PhoP–DNA Complex. *Biochemistry*, *51*(7), 1346-1356. doi: 10.1021/bi201257d

S. Sowa, G. Gelderman, A. N. Leistra, A. Buvanendiran, S. Lipp, A. Pitaktong, T. Romeo, M. Baldea, and L.M. Contreras, “Integrative FourDomics approach profiles the target network of the carbon storage regulatory system,” Nucleic Acids Research Journal, 45(4): 1673-1686 (2017).

Vakulskas, C., Potts, A., Babitzke, P., Ahmer, B., & Romeo, T. (2015). Regulation of Bacterial Virulence by Csr (Rsm) Systems. Microbiology And Molecular Biology Reviews, 79(2), 193-224. doi: 10.1128/mmbr.00052-14

Winfield, M., & Groisman, E. (2003). Role of Nonhost Environments in the Lifestyles of Salmonella and Escherichia coli. *Applied And Environmental Microbiology*, *69*(7), 3687-3694. doi: 10.1128/aem.69.7.3687-3694.2003

Yamamoto, K., Hirao, K., Oshima, T., Aiba, H., Utsumi, R., & Ishihama, A. (2005). Functional Characterization in Vitro of All Two-component Signal Transduction Systems from Escherichia coli. Journal Of Biological Chemistry, 280(2), 1448-1456. doi: 10.1074/jbc.m410104200